Subject: fast shared storage between containers
Posted by wodzpl on Tue, 29 Oct 2013 16:21:17 GMT
View Forum Message <> Reply to Message

hi,

I am quite new in OpenVZ world.
I am doing a research for specific usage scenario of OpenVZ.
For that I need very fast shared storage between VEs.
In my lab I am using dedicated server platform with 200Gb of ram.
CIFS/NFS from fileserver is not so fast and for extra array with many disks to share the load of 70 VEs read/write operations with many small files budget is tight.
Every VE needs to have r/w rights to that storage.
So I experimented with ramfs/tmpfs and mounting it by bind to VEs.
I need +40 GB ramdisk but every VE I would like to have with limited amount of PHYSPAGES=3GB.
When I try with dd filling ramdisk with file grater thatn 4G oomkiller restores some resources and it`s not accepted.

I`ve been experimenting with SHMPAGES and OOMGUARPAGES but without any luck.

First of all I would like to ask you experienced openVZ community is it possible what I am trying to get? [ramdisk 40GB shared over VEs and VEs with limited 3GB of ram]
Secondly maybe you have other ideas how to provide this sort of very fast shared storage.

I appreciate all your help.

Subject: Re: fast shared storage between containers
Posted by wodzpl on Mon, 04 Nov 2013 17:40:50 GMT
View Forum Message <> Reply to Message

hi folks,

So far I`ve tested with success loop devices [files] created on tmpfs on node and then by bind mount shared to VEs.
This solution is half-measure for me because I was managed to get ~45% slower transfer rate when compared to direct tmpfs mount bind to VE. I assume its too big emulation overhead.

Tmpfs in opposite to ramfs lives in swap and page caches.
VSWAP is located in physical ram.

So I set vswap in VE to size of my ramdisk and then do tests [SizeOfVSWAP=SizeOfTmpfs].
This would work great as planned etc vswap is emulating real swapping to hdd by slowing down my very fast ram sticks ;]
I can image a reason for that feature while using openVZ for vps hostings etc.

But in my case I would like to disable it.

How can I accomplish that?

Subject: Re: fast shared storage between containers
Posted by pavel.odintsov on Sun, 10 Nov 2013 22:51:28 GMT
View Forum Message <> Reply to Message

Hello, wodzpl!

You have so hard challenge:) AFAIK you can create in-memory disk image, mount it via loop driver as block device to hardware node and provide this block device to all containers using vzctl --device. But you must use cluster filesystem and I worry about it support in openvz.
